

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



AMR Vision System for Perception, Job Detection and Identification in Manufacturing

Sarbari Datta and Ranjit Ray

*Robotics and Automation Group, Central Mechanical Engineering Research Institute
India*

1. Introduction

Autonomous mobile robots are becoming an integral part of flexible manufacturing system especially for material transport, cleaning and assembly purpose. The advantage of this type of robots is that the existing manufacturing environment need not be altered or modified as in case of conventional AGVs where permanent cable layouts or markers are required for navigation. These robots are also used extensively for survey, inspection, surveillance, bomb and mine disposal, underwater inspection and space robotics. For autonomous navigation, proprioceptive and exteroceptive sensors are mounted on these mobile robots. As proprioceptive sensors measure the kinematic states of the robot, they accrue error over time and they are supplemented by exteroceptive sensors like ultrasonic and laser range finders, camera and global positioning systems that provide knowledge of its local environment which the robot subsequently uses to navigate. Here we describe the vision system of first indigenous autonomous mobile robot, AMR, with manipulator for environment perception during navigation and for job detection and identification required for material handling in a manufacturing environment.

1.1 Autonomous Mobile Robot System (AMR)

The ultimate goal for research on autonomous navigation of mobile robot is to endow these robots with some practical intelligence so that they can relieve or replace the human operators of tedious and repetitive tasks and for this reason manufacturing is one area where mobile robots are becoming a necessity.

Among on-going research on autonomous mobile robots for applications related to manufacturing, University of Massachusetts Amherst is developing a mobile robot with a comprehensive suite of sensors that includes LRF and vision along with a dexterous manipulator, as mobility extends the workspace of the manipulator, posing new challenges by permitting the robot to operate in unstructured environments (Katz et al., 2006). Bundeswehr University Munich is developing vision-guided intelligent robots for automated manufacturing, materials handling and services, where vision guided mobile robots *ATHENE I* and *II* navigates in structured environments based on the recognition of its current situation and a calibration-free manipulator handles various objects using an stereo-vision system (Bischoff & Graefe, 1998).

In this country, mobile robots are being developed in some research institutes in collaboration with academic institutes and private sectors. One such mobile robot is SmartNav built by Zenn Systems, Ahmedabad in collaboration with IIT, Kanpur and BARC (Sen et al., 2004). Our mobile robot with manipulator, AMR, is especially tailored for material handling and transport in a manufacturing environment. The vehicle navigates autonomously and transports jobs and tools from one workstation to another workstation. Figure 1 shows the AMR with all the mounted sensors. Among the sensors, a stereovision camera is mounted in front of AMR for environment perception. Another CMOS camera mounted on the wrist of the manipulator is used for material detection and identification required for pick and place operation. Laser and sonar range finders are used for localization through map building and for obstacle avoidance respectively during navigation (Datta et al., 2006). AMR stands on a distributed architecture for performing various tasks without any perceptible delay and for safeguarding the total system against major failure that may occur when the total burden rests on a single point of operation (Datta et al., 2007).

2. AMR Perception

2.1 Prior Art

Color image transmission from the robot while navigating in robot workspace has become very important in the field of mobile robotics, not only for localization by feature identification but also for monitoring of the robot environment through reconstructed images at multiple points in the robot work area. The robot work area can be a huge shop floor or a warehouse encompassing an area of about 200meters, for effective control from a remote host through a single WLAN Access Point.



Figure 1. AMR

Traditional transfer of bitmap images is quite cumbersome. A large image transferred over the Ethernet can take several seconds. To alleviate this problem, progressive image transmission scheme is used where image fidelity, taking advantage of such popular image file formats as JPEG (Joint Photographic Experts Group) and GIF (Graphics Interchange Format), is gradually built up so that the viewer can see an approximated image in its whole without the need to wait for all the data to be received (Tong & Zhang, 1998). But gradual building up of an image in a constantly changing environment becomes a hindrance for mobile robot perception, as high-speed image transmission is absolutely necessary while navigating, to capture the changing scenario.

Similarly, several methods exist for reconstruction of transmitted data. Two approaches that provide robust image transmission through reconstruction are decoder-based adaptive reconstruction and reconstruction-optimized source coding (Hemami, 1995).

Among decoder-based adaptive reconstruction, Smoothing Criterion Reconstruction (SCR), an adaptive algorithm, is designed to exploit the characteristics of the compressed visual information, which reconstructs the lost information of the image using image characteristics such as spatial and temporal correlation (Hemami & Meng, 1995). As such, SCR generally requires extensive computation power, which thwarts the purpose of online viewing of robot environment through fast reconstruction during navigation. Another approach, Vector Quantized Linear Interpolation (VQLI) (Hemami & Gray, 1994) provides reconstruction of equivalent visual quality with less than 10% transmission overhead. Vector Quantization (VQ) is used at the encoder to set appropriate weights for image compression which is decoded for reconstruction. This approach provides reconstruction capabilities without the extensive computational burden as in previous case, but restricts coding of the image for transmission to a proprietary format.

In reconstruction-optimized source coding, a block-based source coding technique Lapped Orthogonal Transform (LOT) is designed to maximize the reconstruction performance (Hemami, 1996). Mean-reconstruction, in which a missing coefficient block is replaced with the average of its available neighbors, is selected and a reconstruction criterion is defined for equal distribution of reconstruction errors across all transform coefficients and a family of LOT is then designed to meet the reconstruction. The overall performance can be gauged by considering both the coding gain and the reconstruction gain. Although the reconstruction-optimized LOT family provides excellent reconstruction capability, but any kind of matrix manipulation required is inconvenient for instant viewing of robot environment through fast reconstruction.

2.2 AMR Image Transmission Network

Reliable transmission and reception of images is imperative for mobile robot perception. Most transmission schemes take advantage of popular JPEG or GIF image format as responsiveness gained from rapid image transmission is more important than perfect image fidelity. Robustness is therefore vital for rapid image transmission and reconstruction in a mobile robot network. Hence, we also take advantage of the most popular and widely supported JPEG image file format (Wallace, 1991) (Schafer, 2001) for transmitting full color images frame by frame from AMR to multiple clients, set at different monitoring points within AMR work area in a manufacturing environment and for reconstructing these images for viewing almost without any perceptible delay.

2.2.2 Methodology for Image Transmission and Reconstruction

A. Image Transmission

High-resolution color images are transmitted over the network using socket communication. Multi-threading is used so that the AMR server can cater to multiple clients. Microsoft Foundation Class (MFC) **CSocket** is used, as this class is highly useful for client/server model communication. **CSocket**, derived from the base class **CObject**, uses serialization protocol to pass data to and from a socket object via a **CArchive** object. An intermediate class **CSocketFile**, also derived from base class **CObject**, is required, as the **CArchive** object attaches to an object of class **CSocketFile** for sending or receiving data.

In our case, for the server side a **CServerSocket** object, inherited from **CSocket**, is created. One **CArchive** object is created for sending data and one for receiving data from the clients, which is associated with **CSocketFile** object in the **CArchive** constructor. The server socket is set in listening mode and on accepting a client, it creates a new object of the class **CListenSocket**.

After an image is grabbed, the image data is assembled which is written to the listening socket for sending to the client through **CArchive** object.

Similarly, for images grabbed from stereovision camera, the images are converted from raster format to JPEG format for ease of transmission using standard technique for image compression. Image data is then serialized using **CArchive** class and written to **CSocket** using **CSocketFile** for transmission to clients.

B. Image Reconstruction

The image data is sent from the host to the clients over WLAN and Ethernet. Data is received on the client side through **CArchive** object, which in turn accesses **CClientSocket** inherited from **CSocket** via **CSocketFile**.

Once the client receives the image data, it reconstructs the image using the COM (Component Object Model) class, **IPicture**. **IPicture** manages a picture object and its properties. Picture objects provide a language-neutral abstraction for bitmaps, icons, and metafiles. A class **CPicture** is created which holds an ATL (Active Template Library) smart pointer **CComQIPtr** to the **IPicture** interface. Class **CPicture** encapsulates only those methods needed for displaying the images. The image data received by the client is loaded in the memory as a **CMemFile** object using **CArchive** class. **CMemFile** is the **CFile**-derived class that supports memory files.

The image is loaded as a stream using COM class **IStream**, which calls **OleLoadPicture** to load the image in the memory. Finally, **Render** is called at a specified offset which instantiates **IPicture** method for rendering the image onto specified device context. The block diagram in Figure 3 shows the total image transfer scheme described here.

3. Object Detection and Identification for Material Handling

AMR material handling system consists of Intelitek's 5 DOF Scrobot-ER-4u manipulator with a CMOS camera, mounted on the wrist of the manipulator as shown in Figure 4. Once the AMR navigates its way to the target workstation, the manipulator routine is invoked. Figure 5 shows the overall AMR software architecture. SBC2, acting as the server, turns on the manipulator control box, which in turn activates the manipulator. As the manipulator moves over the worktable, the camera scans the worktable and when it detects a job, using template matching identifies the desired job or tool. Adaptive thresholding is used for dynamic image segmentation. Using the gray level distribution of an image, the

neighborhood around the highest peak of the histogram is chosen as the threshold region. A novel variation of Otsu’s method (Otsu, 1979) is proposed for faster online processing, which chooses the optimal thresholds by maximizing the between-class variance with an heuristic search method for adaptive thresholding. Once the job parameters are calculated, the gripper picks the job and puts it on the platform. In this way, the jobs are stacked on the vehicle and are transported to the next workstation where they are unloaded using reverse operation and AMR continues with its next mission.

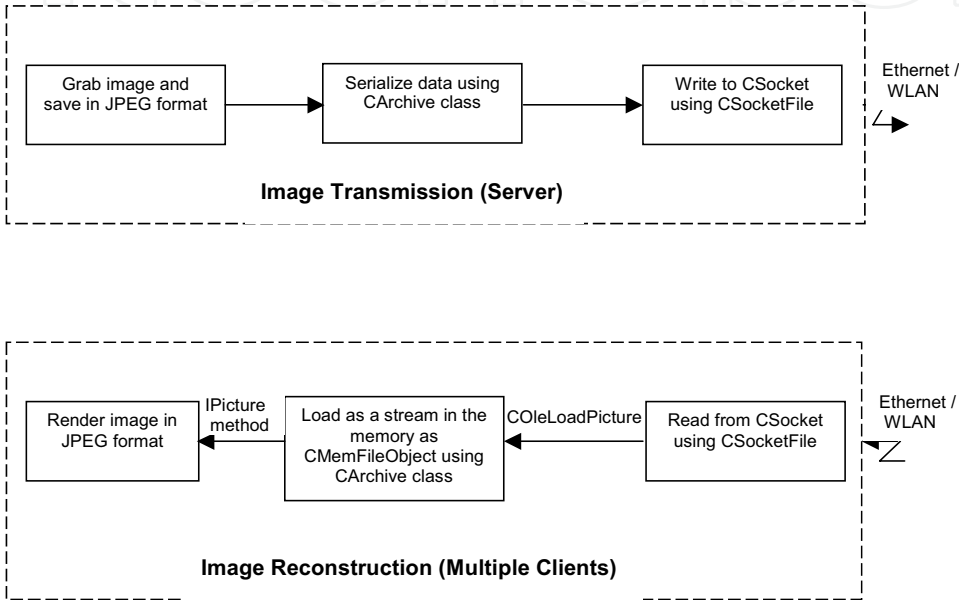


Figure 3. Block diagram of total image transfer scheme



Figure 4. Manipulator setup for material handling using template matching

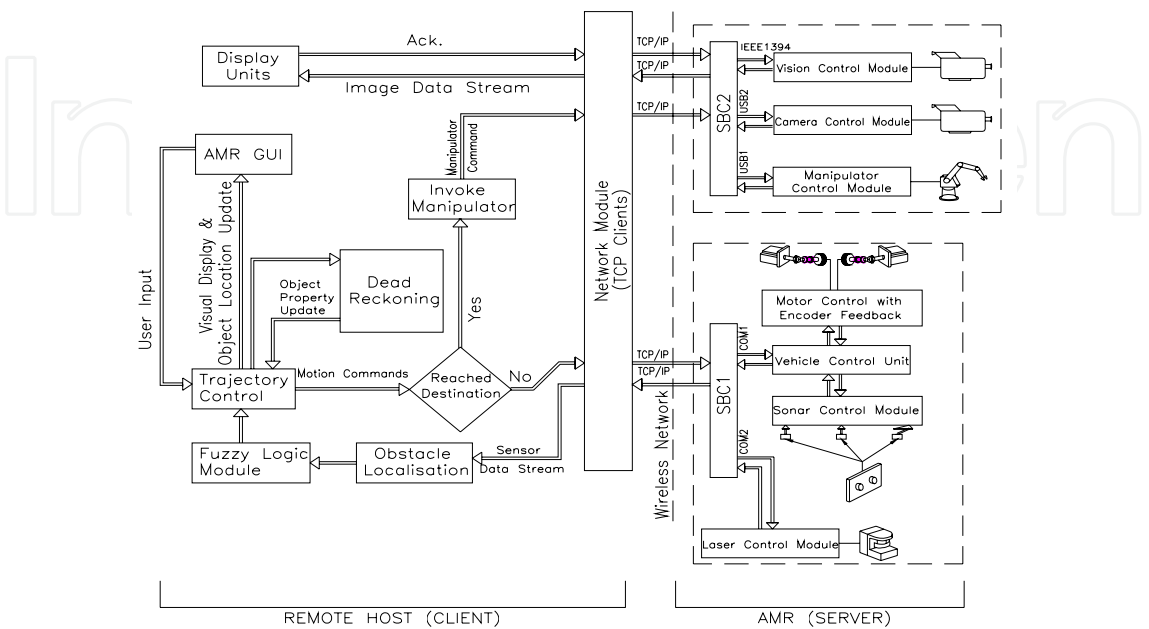


Figure 5. AMR software architecture

3.1 Correlation based Online Template Matching

Template matching is a proven process for classifying unknown samples by comparing them or matching them to known prototypes or templates. The matching process involves (1) moving the template within a search area, the search area may be a sub-image within a larger image or a whole image area (2) at each template location, computing the similarity between the template and the image area over which the template is positioned, and (3) determining the position where a similarity measure is obtained.

The measure of match, $M(f,g)$, represents the degree of similarity between two digital images, $\{f\}$ and $\{g\}$. Correlation measure is one of the few methods to gauge the “measure of match”. Other methods include inter-pixel distance measure where measures of match are based on the pixel-by-pixel intensity differences between two images $\{f\}$ and $\{g\}$; sequential similarity detection algorithms (Barnea and Silverman, 1972), which proposes a more efficient alternative to correlation measures for template matching where the measure of match is based indirectly on an error measure for corresponding pixels in $\{f\}$ and $\{g\}$, the images under comparison at any stage during registration process; and sign change criteria (Venot et al., 1984) for registration of dissimilar images where if we take the pixels in the difference image of two images $\{f\}$ and $\{g\}$ which differ only by additive noise with zero mean and a symmetry density function, i.e.

$$d_{ij} = f_{ij} - g_{ij}$$

In row-column order, there will be many sign changes between adjacent d_{ij} and images which have differences significantly greater than the mean of the noise will not produce

many sign changes between adjacent d_{ij} , which is the motivation behind the use of sign change criteria, a basis for measures of similarity between images.

But the most prevalent method for measure of similarity is the correlation measure. The correlation between the template and the image window has been used as a measure of similarity in template matching and image registration since the 1970s (Rosenfeld, 1969).

For digitized images $\{f\}$ and $\{g\}$ of size A , the normalized correlation coefficient ($corr$) between $\{f\}$ and $\{g\}$ is defined as

$$corr(f, g) = \frac{E[(f - E[f]) \cdot (g - E[g])]}{sd[f] \cdot sd[g]}$$

which is usually simplified to

$$corr(f, g) = \frac{E[f \cdot g] - E[f] \cdot E[g]}{sd[f] \cdot sd[g]}$$

where $E[x]$ is the expected value or mean of a data set $\{x\}$ and $sd[x]$ is the standard deviation of $\{x\}$. The correlation coefficient takes a value in the range of -1.0 to $+1.0$, providing a quantitative measure of similarity between two data sets.

Though the advantages of the correlation coefficient approach are its reliability and accuracy however, computing the correlation coefficient is extremely expensive. The calculation of correlation coefficients for every possible search point during template matching is extremely time consuming. Thus a search method with both high speed and accuracy is required in making the correlation coefficient method computationally reasonable.

Among fast template matching techniques, bounded partial correlation (BPC), based on the normalized cross-correlation (NCC) is used for finding global distortion minimum or correlation maximum (Stefano & Mattoccia, 2003). It is an extension of successive elimination algorithm (SEA) (Li & Salari, 1995) (Wang & Mersereau, 1999) and partial distortion elimination (PDE) (Bei and Gray, 1985), which allow for notably speeding up the computation required by an exhaustive-search template-matching process. Since BPC is a data dependent optimization technique, the computational benefit depends on the image, the template, the position of the template within the image, the correlation threshold, as well as on whether or not one deploys information concerning the expected matching position.

(Yoshimura & Kanade, 1994) suggest using multi-resolution eigenimages for fast template matching based on normalized correlation. This method allows to accurately detect both location and orientation of the object in a scene at faster rate than applying conventional template matching to the rotated object.

Another existing template matching technique is the use of sum-of-squared-differences (SSD) measure to determine the best match. Unfortunately, this measure is sensitive to outliers and is not robust to variations in the template, such as those that occur at occluding boundaries in the image. To compensate for these drawbacks techniques such as subpixel localization, uncertainty estimation and optimal feature is used for robust measure (Olson, 2000).

Another traditional technique for template matching using cross-correlation and an exhaustive search is Fast Fourier transform (FFT) operations which can be used to calculate the cross-correlation surface (Anuta, 1970). In order to use an FFT, the image dimensions (N) must be powers of 2. Therefore it is necessary to pad the template with zeroes in order to make it the same size as the image.

(Vanderbrug & Rosenfeld, 1977) using sum of the absolute valued differences (SAVD) and (Goshtasby et al., 1984) using cross-correlation describe two-stage template matching for reducing the computation required of template matching. This two stage template matching is refined to coarse-fine template matching (Rosenfeld & Vanderbrug, 1977) where a low - resolution template is applied in the first stage, followed by the full resolution template where the match threshold is exceeded. Another class of fast search algorithms is three-step search (Jain, 1981), which is widely used in motion estimation for digital video compression and processing. In the first search step, a search step size of 4 pixels is used. Once an optimal point is found, the step size is reduced to 2 pixels to evaluate the neighborhood of this previously determined optimal point to choose the next search point. In the third step, all the neighboring points of second search point are evaluated to find the final best-matched point. Certainly, this fast search method can speed up the search process, but mismatches or suboptimal matches can occur.

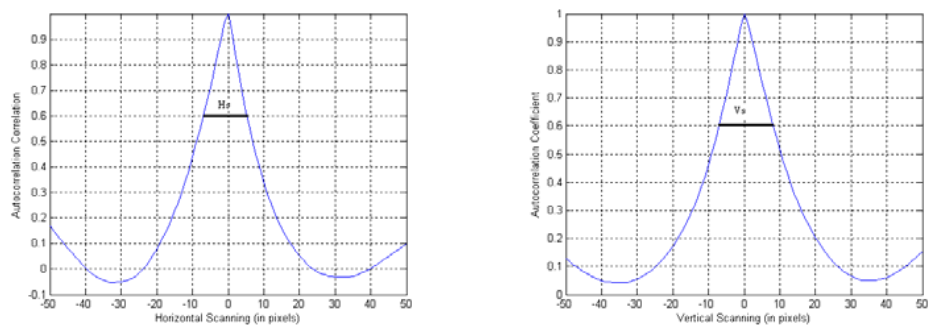


Figure 7. Selection of horizontal and vertical search steps

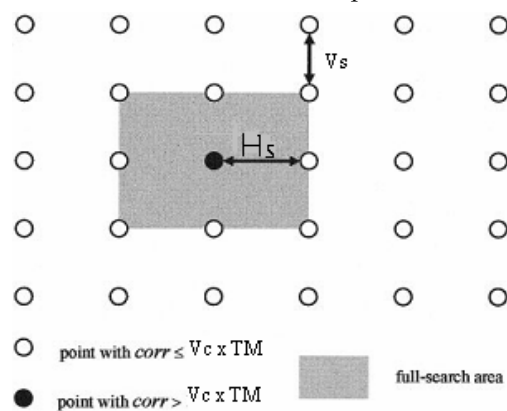


Figure 8. CAPS search lattice

In our AMR vision system application for object detection and identification, after the robot navigates its way to the target workstation, the CMOS camera, mounted on the wrist of a 5 DOF articulated manipulator, identifies pre-defined jobs for pick and place operation using Correlation-based Adaptive Predictive Search (CAPS) method (Shijun et al., 2003), which is based on coarse-fine search method. Using predetermined characteristics computed from its

autocorrelation, CAPS method justifiably selects a set of horizontal and vertical search steps rather than consecutive point-to-point search for faster job detection as shown in Figure 7. For a particular cut-off coefficient V_C from the autocorrelation graph in Figure 7, the horizontal and vertical widths, H_s and V_s , are chosen as step sizes for coarse search. Figure 8 shows the CAPS search lattice with CAPS horizontal and vertical step sizes H_s and V_s . In our case, H_s and V_s are determined for $V_C = 0.6$, which yields satisfactory result.

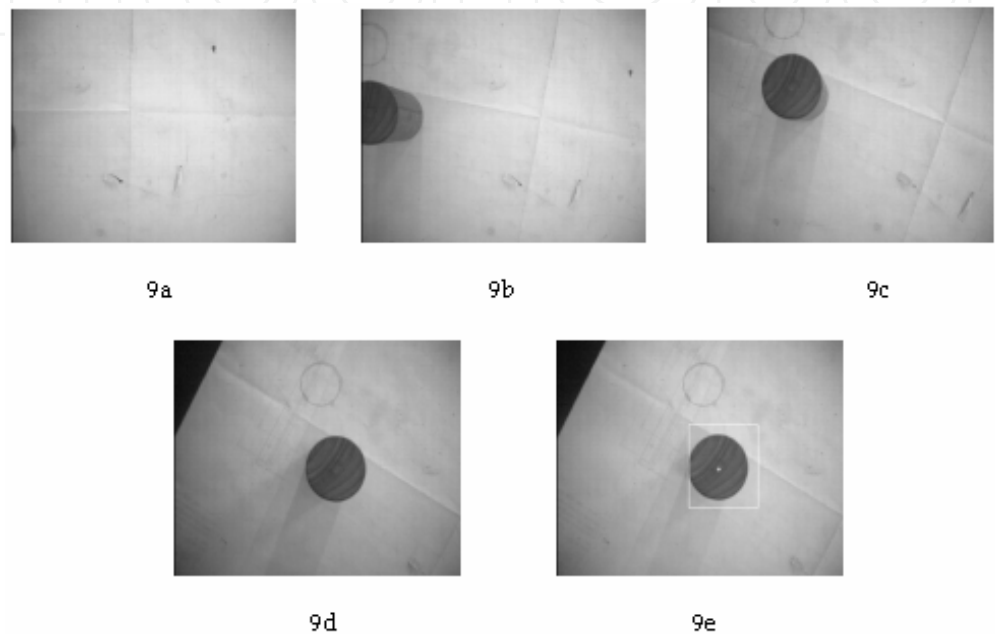


Figure 9. Application of CAPS method on 280x352 image

For our implementation, as the arm moves over the worktable, the camera scans the worktable. As it approaches the job, with each scan, using coarse search technique through pre-calculated vertical and horizontal steps, the correlation coefficient with respect to the stored template is calculated to find out the tentative pose of job and the pose information is then transferred, first with respect to camera and then with respect to manipulator base. Next the camera is moved to this position. Figure 9 shows the sequence of identifying a job based on CAPS method. Figure 9a, 9b and 9c show the result at the end of each coarse search. When the correlation coefficient is greater than matching threshold value T_M , based on the statistics of the template, through fine search technique actual pose of the job is calculated as shown in Figure 9d. In our case, using environment conditions which includes illumination of the work area, the value of $T_M = 0.8$ is found suitable. Figure 9e finally identifies the actual location of the image before thresholding for parametric calculation. Figure 10 gives the block diagram of the CAPS algorithm for template matching.

3.2 Image Thresholding for Proper Gripping

Thresholding is an important and most commonly used technique for image segmentation that tries to identify and extract the image of an object from its background on the basis of

the distribution of grey levels in the captured image. Thresholding techniques can be categorised into two classes: global thresholding and local (adaptive) thresholding. In global thresholding, a single threshold value is used to separate the foreground and the background of an image. It is attractive because it is simple and is sufficient in a fixed, structured environment. However, in case of AMR navigating its way from one workstation to another, due to uneven illumination, local thresholding is more appropriate for segregating the image from the background for proper gripping of the object through parametric calculations using Freeman chain coding (Freeman, 1961).

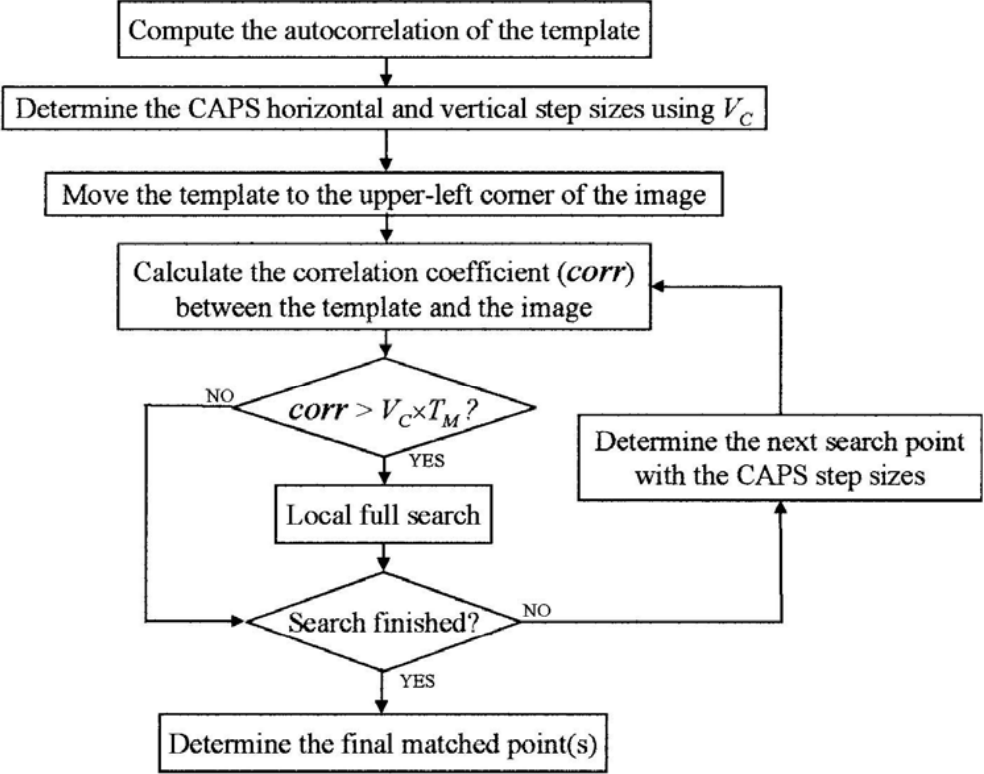


Figure 10. Block diagram of CAPS method

Over the years many image thresholding techniques have been developed and considerable research continues nowadays (Sahoo et al., 1988). The reason for this longterm, ongoing effort is that none of the methods are capable of optimal performance under all conditions. Thresholding selection techniques can be primarily divided into two groups - bilevel and multilevel thresholding. In an image, if the object is distinct from the background, then the histogram of the grey-level is bimodal. For bilevel thresholding, threshold value is selected that coincides with the valley of the grey-level histogram. Multilevel thresholding is used when the histogram of the greyscale image is multimodal. For real time implementation, most thresholding techniques are based on the statistics of the one-dimensional (1D) histogram of grey levels. Many 1D thresholding methods have been investigated. Among frequently used optimal thresholding methods is entropic

thresholding. (Pun, 1980; Kapur et al., 1985) proposes an approach, which maximizes *a posteriori* entropy to measure the homogeneity of threshold classes, (Sahoo et al., 1997) proposes a variation to this approach through Renyi entropy. However, these methods are computationally intensive, hence time consuming thus not suitable for real time computation. (Sahoo et al., 1988) in their study on global thresholding concluded that Otsu's method was one of the better threshold selection methods for general real world images with regard to uniformity and shape measures. Otsu's method chooses the optimal thresholds by maximizing the between-class variance with an exhaustive search (Otsu, 1979). The drawback of Otsu's method is as the number of classes of an image increase, Otsu's method exceeds the time limit for multilevel thresholding in real time. To overcome this, (Liao et al.,) proposes a modified approach based on heuristic search method for faster multi-level thresholding.

For our AMR, we have selected one dimension bi-level thresholding using a maximum of eighteen-step on-line heuristic search on a gray-scale image based on Otsu's method for determining the proper image threshold.

3.2.1 Eighteen Step Algorithm for on-line thresholding

Defining Otsu's method for image thresholding, an image is a 2D grayscale intensity function containing N pixels with gray levels from 1 to L. The number of pixels with gray level i is denoted f_i , giving a probability of gray level i in an image of

$$p_i = f_i/N \quad (1)$$

In the case of bi-level thresholding of an image, the pixels are divided into two classes, C1 with gray levels [1, ..., t] and C2 with gray levels [t+1, ..., L]. Then, the gray level probability distributions for the two classes are

C₁: $p_1/\omega_1(t), \dots, p_t/\omega_1(t)$ and C₂: $p_{t+1}/\omega_2(t), p_{t+2}/\omega_2(t), \dots, p_L/\omega_2(t)$ where

$$\omega_1(t) = \sum_{i=1}^t p_i \quad (2)$$

and

$$\omega_2(t) = \sum_{i=t+1}^L p_i \quad (3)$$

Also, the means for classes C₁ and C₂ are

$$\mu_1 = \sum_{i=1}^t i \cdot p_i / \omega_1(t) \quad (4)$$

and

$$\mu_2 = \sum_{i=t+1}^L i \cdot p_i / \omega_2(t) \quad (5)$$

Let μ_T be the mean intensity for the whole image. It is easy to show that

$$\mu_1 \cdot \omega_1 + \mu_2 \cdot \omega_2 = \mu_T \quad (6)$$

$$\omega_1 + \omega_2 = 1 \quad (7)$$

Using discriminant analysis, Otsu defined the between-class variance of the thresholded image as

$$\sigma_B^2 = \omega_1 \cdot (\mu_1 - \mu_T)^2 + \omega_2 \cdot (\mu_2 - \mu_T)^2 \quad (8)$$

For bi-level thresholding, Otsu verified that the optimal threshold t^* is chosen so that the between-class variance σ_B^2 is maximized; that is,

$$t^* = \text{Arg Max}_{1 \leq t \leq L} \{ \sigma_B^2(t) \} \quad (9)$$

As an alternate formulation to Otsu's method, using Eqs. (6) and (7), the between-class variance in Eq. (8) of the thresholded image can be rewritten as

$$\sigma_B^2 = \omega_1 \mu_1^2 + \omega_2 \mu_2^2 - 2\mu_T^2 \quad (10)$$

As the last term of Eq. 10 is independent of the choice of the thresholds, the optimal bi-level threshold is chosen by maximizing a modified between-class variance ($\dot{\sigma}_B^2$), defined as

$$\dot{\sigma}_B^2 = \omega_1 \mu_1^2 + \omega_2 \mu_2^2 \quad (11)$$

Hence, Eq. 6 can be written as

$$\mu_T = \sum_{i=1}^L i \cdot p_i \quad (12)$$

From Eqs. 6 & 7, modified between-class variance ($\dot{\sigma}_B^2$) can be written as

$$\dot{\sigma}_B^2 = \omega_1 \mu_1^2 + \frac{(\mu_T - \omega_1 \mu_1)^2}{(1 - \omega_1)} \quad (13)$$

Comparing Eq. 13 with Eq. 8, we find that $\dot{\sigma}_B^2$ value can be directly calculated ignoring the Eq.3, Eq.5 & Eq.6. Again, from the Eq. 9., by OTSU method, optimal bi-level threshold is chosen by maximizing modified between-class variance ($\dot{\sigma}_B^2$) for the gray level from 1 to L.

According to the criteria of both Eq. 9 for σ_B^2 and Eq. 13 for $\dot{\sigma}_B^2$ to find the optimal threshold by Otsu method, the search range for the maximal σ_B^2 and the maximal $\dot{\sigma}_B^2$ is $1 \leq t^* < L$.

This exhaustive search involves (L-1) possible combination, computationally intensive for on-line processing. Using eighteen-step method, we have reduced the computational time by reducing the exhaustive search from (L-1) possible combinations to a maximum of eighteen combinations for detecting the proper threshold. The flowchart in Figure 11 delineates this process.

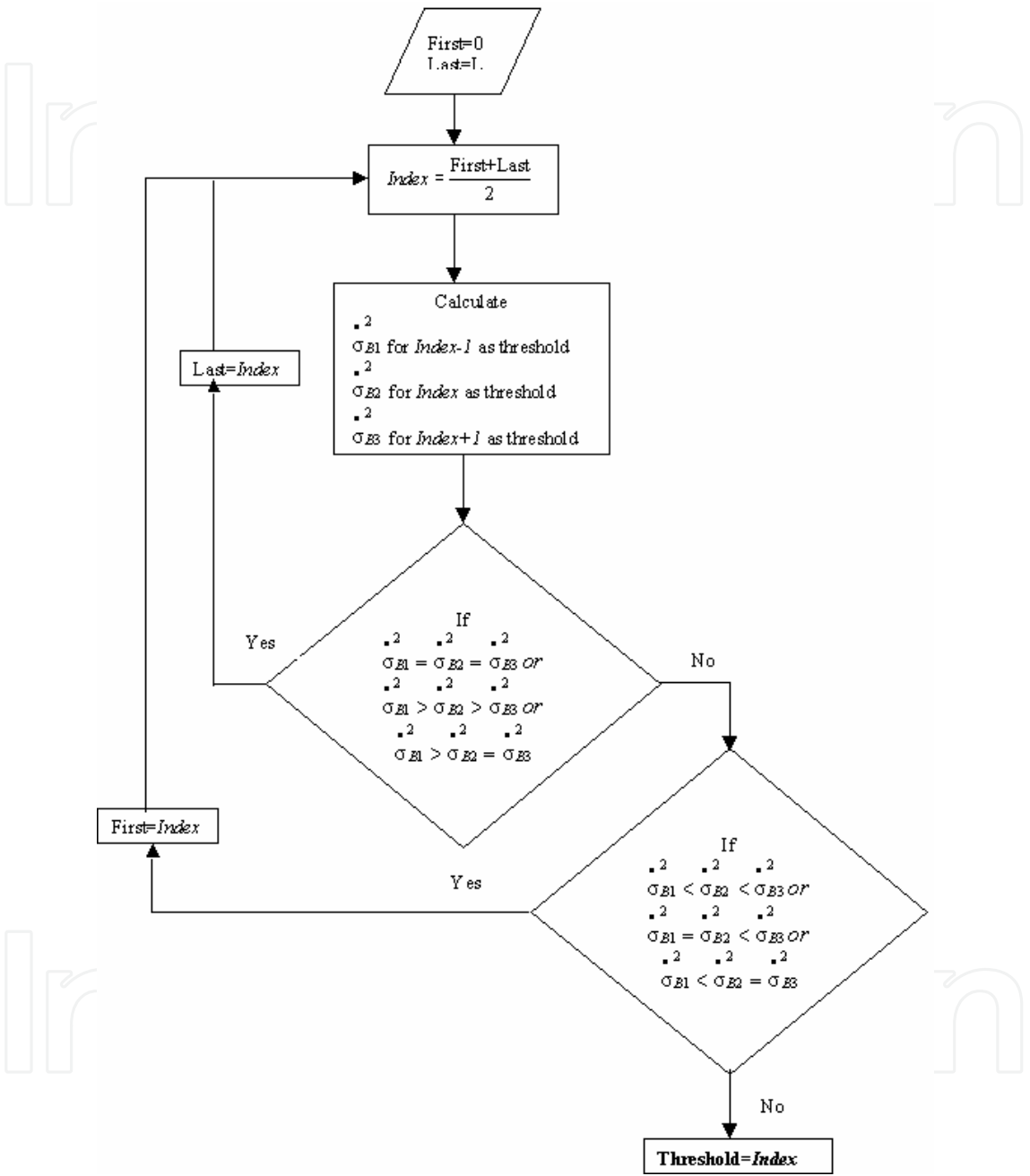


Figure 11. Flowchart of proposed eighteen-step method

4. Results and Discussion

4.1 Analysis of transmission and reconstruction rate for AMR perception

Experimental data of image transmission rate and reconstruction rate is recorded using the 100 Mbps Ethernet and 11Mbps WLAN for mono-vision and stereovision cameras. The data is tabulated, analyzed, graphically illustrated and analyzed in this section.

A. Image transmission and reconstruction using mono-vision camera:

Using the approach discussed above, several processes are tabulated. First, colour images of resolution 768x576 are transmitted frame by frame from a mono-vision camera over 100 Mbps Ethernet and 11 Mbps WLAN and are reconstructed on the client side. The grab rate is around 13 fps, which is hardware dependent, comprising a mono-vision PULNIX-TMC-6DSP camera with a Matrox Meteor-II/Standard frame grabber card. Next the image frame is serialized and is processed as a **CSocket** object. The image data is then written to the **CListenSocket** for transmission. Finally, there is an acknowledgement from the client side as per TCP/IP protocol, before the server is ready for sending the next frame. The process for image transmission is tabulated in Table 1 below:

| Approximate time in milliseconds | | | |
|----------------------------------|--|----------|-----------|
| | Process | Ethernet | WLAN |
| 1. | Assembling an image frame for transmission | 110 | 110 |
| 2. | Acknowledgement from client | 50 | 90-120 |
| 3. | Total process time | 160 | 200 - 230 |

Table 1. Time for transmitting a 768x576 JPEG image frame

Next, the process on the client side is recorded. For fast, uninterrupted display, the image data is reconstructed using COM class, **IPicture**. The image data received by the client is loaded as a stream, using COM class **IStream**, in the memory as a **CMemFile** object using **CArchive** class. The breakup of the total process time for receiving the frame-by-frame image data on the client side through Ethernet and over WLAN is given in Table 2. Reading serialized data for reconstruction varies inversely with the throughput rate of the medium and there is no perceptible difference between transmitting image frames over these two media when it comes to viewing the environment through on-line reconstruction of the scene.

| Approximate time in milliseconds | | | |
|----------------------------------|-------------------------------|----------|---------|
| | Process | Ethernet | WLAN |
| 1. | Reading serialized image data | 280 | 280-380 |
| 2. | Displaying an image frame | 50 | 50 |
| 3. | Total process time | 330 | 330-430 |

Table 2. Time for displaying a 768x576 JPEG image frame

The nature of image transmission over Ethernet and WLAN is graphically illustrated in Figure 12 and Figure 13.

As the throughput rate of Ethernet is higher than that of WLAN, barring few aberrations, total process time for transmitting an image frame over Ethernet is around 160ms while transmitting the same frame over WLAN takes between 200 ms and 230 ms, as given in Table 1. Figure 13 shows that transmission over WLAN is more prone to environmental noise, common in a manufacturing environment. Unlike transmission over Ethernet where a steady rate is maintained, variable transmission rate over WLAN does not hamper the

scene-by-scene update of the environment when it comes to viewing through online reconstruction of the scene.

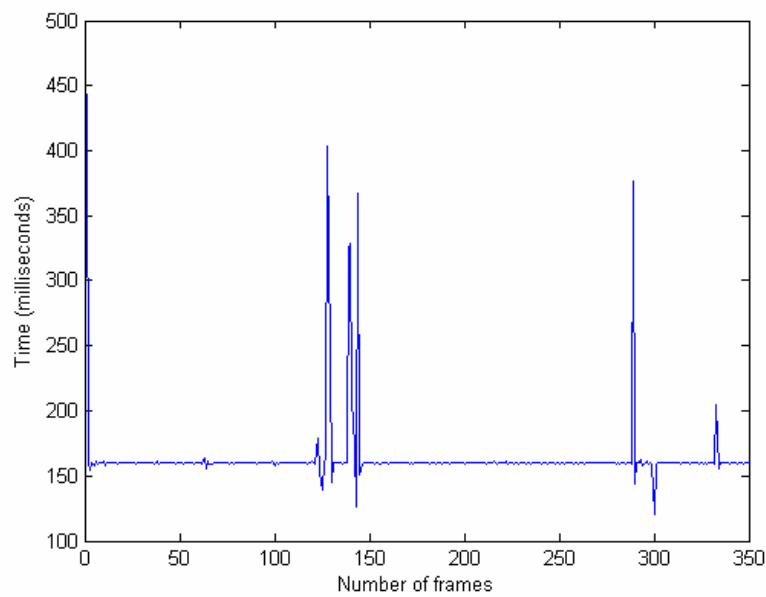


Figure 12. Time for sending a 768x576 image frame over Ethernet

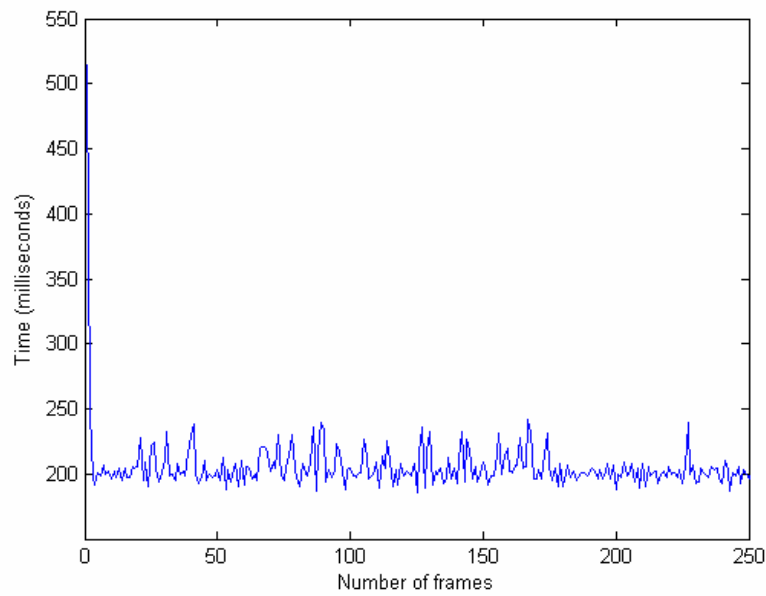


Figure 13. Time for sending a 768x576 image frame over WLAN

B. Image transmission and reconstruction using stereovision camera:

Using the same procedure as mentioned in the above section, stereo images of resolution 1024x768 are converted from raster format i.e. from PPM format to JPEG format. From grabbing the image in PPM format to converting it to JPEG format takes around 1.04 seconds as shown in Figure 14. Hence, with other parameters remaining the same, as described in the above section, total process time for transmitting an image frame over the Ethernet is less than 1.2 seconds and it hovers around 1.3 seconds over WLAN.

4.2 Result of CAPS based template matching

We have implemented CAPS method for online template matching on a 3.6GHz Pentium IV computer running on Microsoft Windows XP platform. For a 76x82 template and the 280x352 image, locating the template using full search took 55.485 seconds whereas using CAPS method with $V_C = 0.6$ and $T_M = 0.8$ took few milliseconds for coarse search and little more than a second for fine search. Table 3 gives a breakup of search method along with search time for the job shown in Figure 9.

4.3 Result of proposed eighteen-step thresholding

For evaluating our proposed eighteen-step algorithm for thresholding, we have considered four conventional gray images of F16 jet, Baboon, Lena and Peppers of the size 128x128 pixels as shown in Figure 15. Otsu's method, as given in Eq. 8 and Eq. 9, and our proposed algorithm stated in Eq. 13 are implemented in MatLab Version 7.0.1 on a 3.60 GHz Pentium IV computer with Microsoft Windows XP operating system.

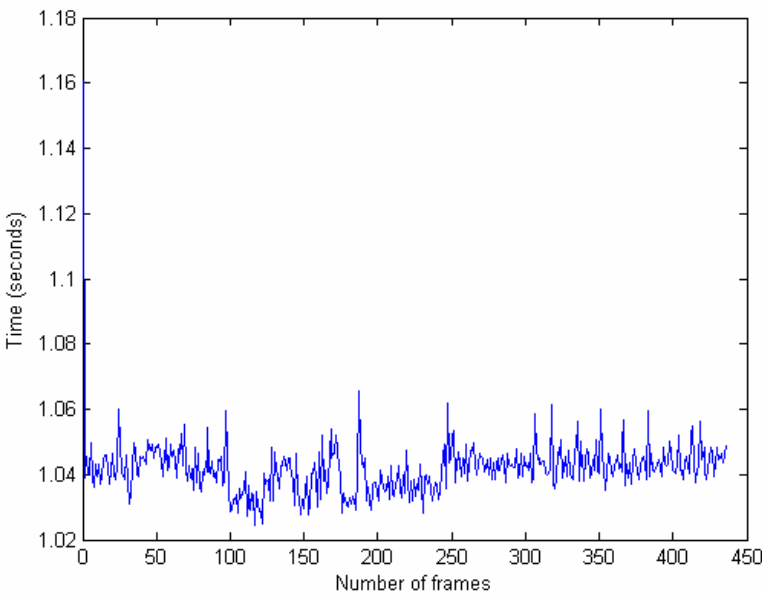


Figure 14. Time for grabbing and converting a 1024 x 768 raster image to JPEG

| | |
|-------------------|----------------|
| Coarse Search | |
| Frame | Time (seconds) |
| Frame 1 (Fig. 9a) | 0.0156 |
| Frame 2 (Fig. 9b) | 0.0156 |
| Frame 3 (Fig. 9c) | 0.125 |
| Fine Search | |
| Frame 4 (Fig. 9d) | 1.422 |

Table 3. Computation time for template matching using CAPS method

Table 4 gives the comparative result between Otsu’s method and eighteen-step method for bi-level thresholding on these four test images. Figure 16 shows the plot of modified between class variance σ_B^2 against corresponding gray level values required for determining the proper bi-level threshold. The peak for each image sets the threshold for that image. Finally, Figure 17 shows the thresholded test images.

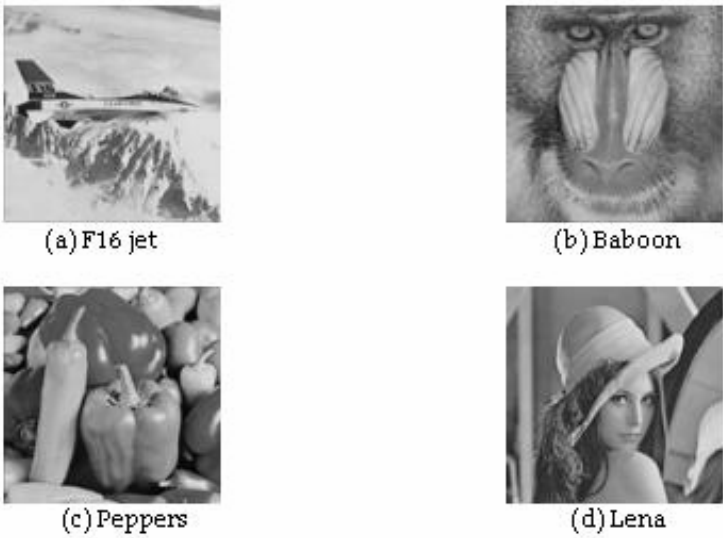


Figure 15. Test images

| Images | Computation time (milliseconds) | | Bi-Level Threshold |
|---------|---------------------------------|----------------------|--------------------|
| | Otsu’s method | Eighteen-Step method | |
| F16Jet | 140 | 16 | 155 |
| Baboon | 141 | 16 | 130 |
| Peppers | 141 | 16 | 121 |
| Lenna | 219 | 15 | 118 |

Table 4. Evaluation of Eighteen-Step method on test images for bi-level thresholding

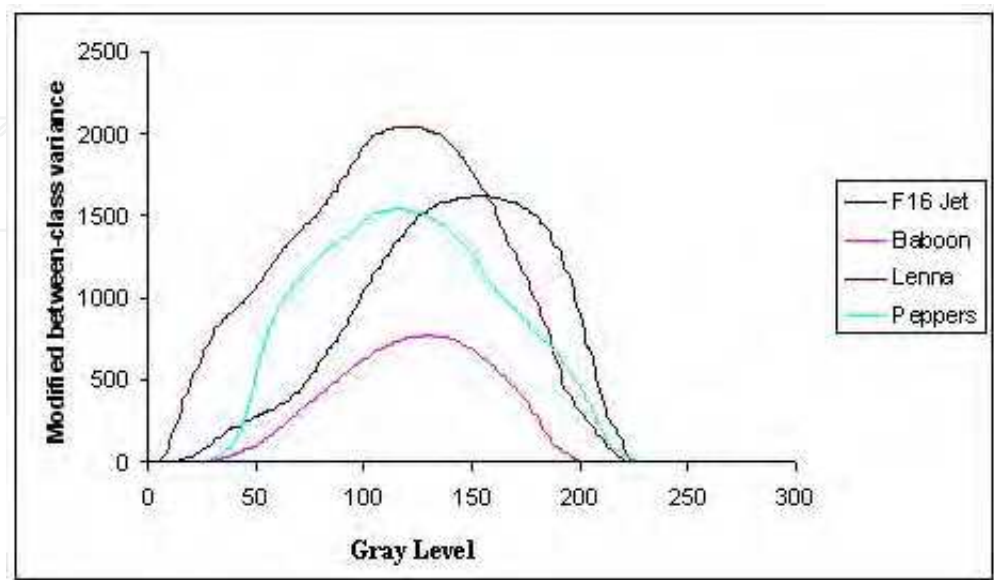


Figure 16. Plot of modified between-class variance against the corresponding gray level value

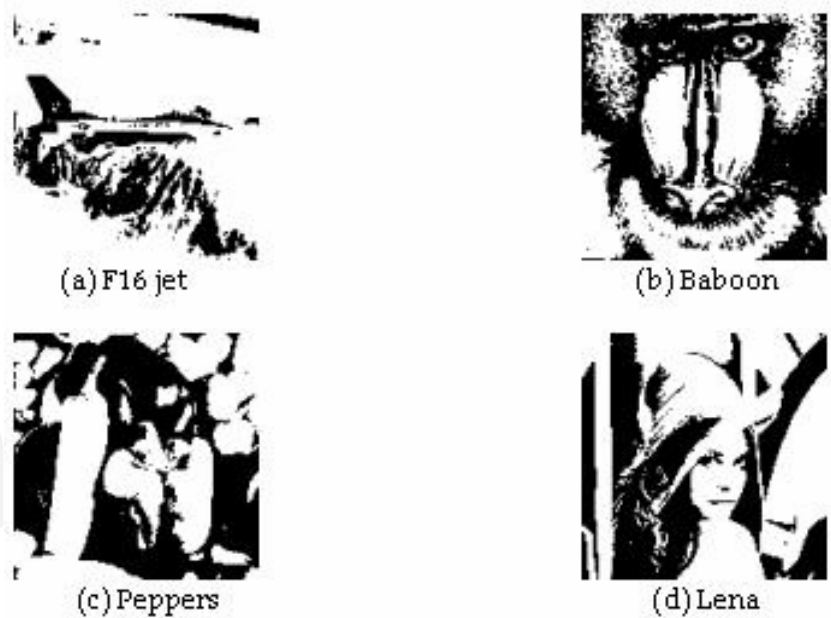


Figure 17. Thresholded images

7. Conclusion

Vision is becoming an integral part of robotic systems not only for navigation but also for job identification for material handling as camera is the only sensor that imparts a feel of spatial sensing through 3-D sensing that is lacking in other sensors like laser or ultrasonic range finders. In AMR, vision plays an integral role in all aspects. It plays a pivotal role for mobile robot perception of environment while navigating, through rapid transmission of images from mobile robot to remote viewers. Moreover, vision also plays a crucial role in online job identification for job handling. Though for all these above tasks, the sheer volume of information to be processed online becomes a hindrance, as it was in the past, but we have shown that by coming up with novel concepts based on existing knowledge and ideas and with continuing advancement in computer architecture, especially with powerful modern processors available today, we can not only overcome these difficulties but use its unique feature to our advantage.

8. Acknowledgement

The authors would like to thank the members of the project team for their help and support in developing an “Autonomous Mobile Robot for Manufacturing Environment” under the aegis of Council of Scientific and Industrial Research (CSIR, India) network project on Advanced Manufacturing Technology. The authors would like to express their sincere gratitude to the Director of the Institute for his assent in publishing this work.

9. References

- Anuta, P. E. (1970). Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 4, October 1970, pp. 353-368, ISSN: 0359-4237
- Barnea, D.I. & Silverman, H.F. (1972). A class of algorithms for fast digital image registration, *IEEE Trans. on Computers*, Vol. C-21, February 1972, pp. 179-86
- Bei, C.D. & Gray, R.M. (1985). An improvement of the minimum distortion encoding algorithm for vector quantization, *IEEE Trans. on Communications*, Vol. 33, No. 10, October 1985, pp. 1132-33, ISSN: 0096-2244
- Bischoff, R. & Volker, G. (1998). Vision-guided intelligent robots for automating manufacturing, material handling and services, *WESIC'98 Workshop on European Scientific and Industrial Collaboration on Promoting Advanced Technologies in Manufacturing*, Girona, June 1998
- Datta, S.; Ray, R. & Banerji, D. (2007). Development of autonomous mobile robot with manipulator for manufacturing environment, accepted for publication in *International Journal of Advanced Manufacturing Technology*, Springer Publication
- Datta, S.; Banerji, D. & Mukherjee, R. (2006). Mobile robot localization with map building and obstacle avoidance for indoor navigation, *Proc. IEEE International Conference on Industrial Technology*, Vol. 3, pp. 2535-2540, ISBN 1-4244-0726-5, December 15-17 2006, Mumbai, India
- Freeman, H. (1961). On the encoding of arbitrary geometric configurations, *IRE Trans. Electronic Computers*, Vol. EC-10, pp. 260-268, June 1961

- Goshatby, A.; Gage, S.H. & Batholic, J.F. (1984). A two-stage cross correlation approach to template matching, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, pp. 374-78, May 1984
- Hemami, S.S. (1996). Reconstruction-optimized lapped orthogonal transforms for robust image transmission, *IEEE Trans. On Circuits and Systems for Video Technology*, Vol. 6, No. 2, April 1996, pp. 168-181, ISSN: 1051-8215
- Hemami, S.S. (1995). Digital image coding for robust multimedia transmission, *Symposium on Multimedia Communications & Video Coding*, New York, October, 1995
- Hemami, S.S. & Meng, T.H.Y. (1995). Transform-coded image reconstruction exploiting interblock correlation, *IEEE Trans. on Image Processing*, Vol. 4, No. 7, July 1995, pp. 1023-27, ISSN: 1057-7149
- Hemami, S.S. & Gray, R.M. (1994). Image reconstruction using using vector quantized linear interpolation, *Proc. ICASSP '94*, Adelaide, Australia, April 1994, Vol. 5, pp. 629-32
- Jain, J.R. & Jain, A.K. (1981). Displacement measurement and its application in interframe image coding, *IEEE Trans on Communication*, Vol. 29, No.12, December 1981, pp. 1799-1808, ISSN:0096-2244
- Kapur, J.N.; Sahoo, P.K. & Wong, A.K.C. (1985). A new method for gray-level picture thresholding using the entropy of the histogram, *Computer Vision Graphics Image Processing*, Vol. 29, 1985, pp. 273-285
- Katz, D. et al. (2006). The Umass Mobile Manipulator UMan: An experimental platform for autonomous mobile manipulator, *Workshop on Manipulator for Human Environment at Robotics: Science and Systems*, Philadelphia, USA, August 2006
- Li, W. & Salari, E. (1995). Successive elimination algorithm for motion estimation, *IEEE Trans. on Image Processing*, Vol. 4 No. 1, Jan. 1995, pp. 105-107, ISSN:1057-7149
- Liao, P.; Chen, T. & Chung, P. (2001). A fast algorithm for multilevel thresholding, *Journal of Information Science and Engineering*, 17, pp. 713-727, 2001
- Olson, C. F. (2000). Maximum-likelihood template matching, *Proc. IEEE Conference on CVPR*, Vol. 2, pp. 52-57, 13-15 June 2000, Hilton Head, South Carolina, USA
- Otsu, N. (1979). A threshold selection method from gray-level histogram, *IEEE Trans. System Man Cybernetics*, Vol. 9, March 1979, pp. 62-66
- Pun, T. (1980). A new method for gray-level picture thresholding using the entropy of the histogram, *Signal Processing*, Vol.2, 1980, pp. 223-237
- Rosenfeld, A. (1969). *Picture processing by computer*, New York: Academic Press, 1969
- Rosenfeld, A. & Vanderburg, G. J. (1977). Coarse-fine template matching, *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 7, 1977, pp. 104-107
- Sahoo, P.K.; Wilkins, C. & Yeager, J. (1997). Threshold selection using Renyi's entropy, *Pattern Recognition*, Vol. 30, No. 1, pp. 71-84, 1997, Elsevier Science Ltd.
- Sahoo, P.K. et al. (1988). A survey of thresholding techniques, *Computer Vision Graphics Image Processing*, Vol. 41, 1988, pp. 223-237
- Schaefer, G. (2001). JPEG image retrieval by simple operators, *CBMI '01*, pp. 207-214, Brescia, Italy, September 19-21, 2001
- Sen, S; Taktawala, P. K. & Pal, P. K. (2004). Development of a range-sensing, indoor, mobile robot with wireless Ethernet connectivity, *Proceedings of the National Conference on Advanced Manufacturing & Robotics*, pp. 3-10, ISBN 81-7764-671-0, CMERI, January 2004, Allied Publishers Pvt. Ltd., Durgapur

- Shijun, S. et al. (2003). Fast template matching using correlation-based adaptive predictive search, *International Journal of Imaging System Technology*, Vol. 13, pp. 169-178, 2003
- Stefano. L.D. & Mattoccia, S. (2003). Fast template matching using bounded partial correlation, *Machine Vision and Applications*, Vol. 13, pp. 213-21, ISSN: 0932-8092, Springer-Verlag, 2003
- Tong, H.F. & Zhang, D. (1998). A new progressive colour image transmission scheme for the World Wide Web, *Computer Networks and ISDN Systems*, 30 (1998) pp. 2059-2064
- Vanderburg, G.J. & Rosenfeld, A. (1977). Two stage template matching, *IEEE Trans. on Computers*, Vol. C-26, pp. 384-93, April 1977
- Venot, A.; Lebruchec, J.F. & Roucayrol, J.C. (1984). A new class of similarity measures for robust image registration, *Computer Vision, Graphics and Image Processing*, Vol. 28, pp. 176-84, 1984
- Wallace, G.K. (1991). The JPEG Still Picture Compression Standard, *Communications of the ACM*, Vol. 34 No.4, pp. 30-44, 1991
- Wang, H. & Mersereau, R. (1999). Fast algorithm for the estimation of motion vectors, *IEEE Trans. on Image Processing*, Vol. 8, No. 3, March 1999, pp. 435-438, ISSN: 1057-7149
- Yoshimura, S. & Kanade, T. (1994). Fast template matching based on the normalized correlation by using multiresolution eigenimages, *Proc. IROS '94*, Vol. 3, pp. 2086-2093, September 12-16 1994, Munich, Germany



Vision Systems: Applications

Edited by Goro Obinata and Ashish Dutta

ISBN 978-3-902613-01-1

Hard cover, 608 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

Computer Vision is the most important key in developing autonomous navigation systems for interaction with the environment. It also leads us to marvel at the functioning of our own vision system. In this book we have collected the latest applications of vision research from around the world. It contains both the conventional research areas like mobile robot navigation and map building, and more recent applications such as, micro vision, etc. The first seven chapters contain the newer applications of vision like micro vision, grasping using vision, behavior based perception, inspection of railways and humanitarian demining. The later chapters deal with applications of vision in mobile robot navigation, camera calibration, object detection in vision search, map building, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Sarbari Datta and Ranjit Ray (2007). AMR Vision System for Perception, Job Detection and Identification in Manufacturing, Vision Systems: Applications, Goro Obinata and Ashish Dutta (Ed.), ISBN: 978-3-902613-01-1, InTech, Available from:
http://www.intechopen.com/books/vision_systems_applications/amr_vision_system_for_perception__job_detection_and_identification_in_manufacturing

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen